



Filling the Gap Between Lower-C1 and Lower-C2 Functions

Aris Daniilidis, Jérôme Malick

► To cite this version:

Aris Daniilidis, Jérôme Malick. Filling the Gap Between Lower-C1 and Lower-C2 Functions. Journal of Convex Analysis, 2005, 12 (2), pp.315-329. hal-00804407

HAL Id: hal-00804407

<https://inria.hal.science/hal-00804407>

Submitted on 25 Mar 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Filling the gap between lower- C^1 and lower- C^2 functions

ARIS DANIILIDIS & JÉRÔME MALICK

Abstract The classes of lower- $C^{1,\alpha}$ functions ($0 < \alpha \leq 1$), that is, functions locally representable as a maximum of a compactly parametrized family of continuously differentiable functions with α -Hölder derivative, are hereby introduced. These classes form a strictly decreasing sequence from the larger class of lower- C^1 towards the smaller class of lower- C^2 functions, and can be analogously characterized via perturbed convex inequalities or via appropriate generalized monotonicity properties of their subdifferentials. Several examples are provided and a complete classification is given.

Key words Maximum function, lower- $C^{1,\alpha}$ function, α -weakly convex function, α -hypomonotone operator.

AMS Subject Classification *Primary* 26B25; *Secondary* 49J52, 47H05.

1 Introduction

Let U be an open subset of \mathbb{R}^n and $k \in \mathbb{N}^*$. A function $f : U \rightarrow \mathbb{R}$ is called lower- C^k (for short, LC^k), if for every $x_0 \in U$ there exist $\delta > 0$, a compact topological space S , and a jointly continuous function $F : B(x_0, \delta) \times S \rightarrow \mathbb{R}$ satisfying

$$f(x) = \max_{s \in S} F(x, s), \quad \text{for all } x \in B(x_0, \delta),$$

and such that all derivatives of F up to order k with respect to x exist and are jointly continuous. It is easily seen that every such function is locally Lipschitz. In particular, LC^k functions provide a robust extension of both convexity and smoothness. For their role in optimization we refer to the survey [8] and to [18]; see also [17] for extensions in Hilbert spaces.

The class of LC^1 functions is first introduced by Spingarn in [22]. In that work, Spingarn shows that these functions are (Mifflin) semi-smooth and Clarke regular, and that are characterized by a generalized monotonicity property of their subgradients, called submonotonicity. Recently, in [5, Corollary 3], it has been pointed out that the class of LC^1 functions coincides with the class of locally Lipschitz approximately convex functions. We recall that a function $f : U \rightarrow \mathbb{R}$ is called *approximately convex* on U if for every $x_0 \in U$ and $\varepsilon > 0$, there exists $\delta > 0$ such that for all $x, y \in B(x_0, \delta)$ and all $t \in [0, 1]$

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) + \varepsilon t(1-t)\|x - y\|. \quad (1)$$

The above notion (introduced in [14], [15]) corresponds to a first order relaxation of convexity and is strongly related to the notion of α -paraconvexity studied in [11], [20]. A more general class – corresponding to the case that the ε of the above definition is always bounded below away from 0 – is recently considered in [16] for functions on the real line: these functions (which are not Clarke regular in general) are characterized by their local decomposability into a sum of a convex and a Lipschitz function. We refer also to [9] and [7] for related notions.

Shortly after Spingarn's work, the (smaller) class of LC^2 functions has been introduced and studied by Rockafellar [18]. In that work the following important results are established:

- for every $k \geq 2$, the class of LC^k functions coincides with the class of LC^2 functions ;
- LC^2 are exactly the locally Lipschitz weakly convex functions.

We recall that a function $f : U \rightarrow \mathbb{R}$ is called *weakly convex* on U if for every $x_0 \in U$, there exist $\sigma > 0$ and $\delta > 0$ such that for all $x, y \in B(x_0, \delta)$ and $t \in (0, 1)$

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) + \sigma t(1-t)\|x - y\|^2. \quad (2)$$

Let us note that LC^2 functions are characterized by the fact that they are locally decomposable into a sum of a convex continuous function and a concave quadratic function (see [23], [18], [10] e.g.). The existence of a similar decomposition for the class of LC^1 functions remains open (see also Remark 12).

Remark 1 (terminology issues) We wish to draw the attention of the reader on some terminology issues: speaking about locally Lipschitz functions, the classes of weakly convex functions [23], of prox-regular (or proximal retract) functions [2] and of prime-lower nice functions [21] all coincide with the class of LC^2 functions. See also [1], [4], [21] and references therein for related topics.

In this paper, we consider the class of lower- $C^{1,\alpha}$ functions (in short, $LC^{1,\alpha}$), where $0 < \alpha \leq 1$. Roughly speaking, these are LC^1 functions of the form $f(x) = \max_{s \in S} F(x, s)$ for which $\nabla_x F(\cdot, s)$ is α -Hölder (see exact definition in Section 2). We shall show that every such function is characterized by the α -hypomonotonicity (Definition 5) of its (Clarke) subdifferential and enjoys an alternative geometrical description as a $(1 + \alpha)$ -order perturbation of convexity (see Theorem 8). In particular, as the notation suggests, for $\alpha = 1$ we recover the class of LC^2 functions (see Remark 9).

2 Prerequisites and definitions

Let $f : U \rightarrow \mathbb{R}$ be a locally Lipschitz function defined in an open subset U of \mathbb{R}^n . For every $x_0 \in U$, the (Clarke) generalized derivative of f at x_0 is defined as follows:

$$f^o(x_0; d) := \limsup_{(y,t) \rightarrow (x_0, 0+)} \frac{f(y + td) - f(y)}{t}, \quad \text{for all } d \in \mathbb{R}^n.$$

It follows (see [3, Proposition 2.1.1], for example) that $d \mapsto f^o(x_0; d)$ is a continuous sublinear functional, so that the Clarke subdifferential $\partial f(x_0)$ of f , that is, the set

$$\partial f(x_0) = \{x^* \in \mathbb{R}^n : f^o(x_0; d) \geq \langle x^*, d \rangle, \forall d \in \mathbb{R}^n\} \quad (3)$$

is nonempty. In particular, the multivalued operator $\partial f : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ given by (3) if $x \in U$ and being empty for $x \in \mathbb{R}^n \setminus U$ is called subdifferential of f . If f is a C^1 function then $\partial f(x) = \{\nabla f(x)\}$, for all $x \in U$. Natural operations in optimization (as for instance taking the maximum of an index family of differentiable functions) often lead to nonsmooth functions, in which case ∂f is used to substitute the derivative. We refer to the classical textbooks [3], [4] and [19] for details and applications to optimization.

In this work we study a particular class of maximum-type locally Lipschitz functions. Let us give the following definition.

Definition 2 (lower- $C^{1,\alpha}$ function) Let U be an open set of \mathbb{R}^n , and $0 < \alpha \leq 1$. A locally Lipschitz function $f : U \rightarrow \mathbb{R}$ is called lower- $C^{1,\alpha}$ at $x_0 \in U$, if there exist a non-empty compact set S , positive constants $\delta, \sigma > 0$ and a continuous function $F : B(x_0, \delta) \times S \rightarrow \mathbb{R}$ which is differentiable with respect to the x -variable, such that

$$f(x) = \max_{s \in S} F(x, s), \quad \text{for all } x \in B(x_0, \delta),$$

where $\nabla_x F(x, s)$ is (jointly) continuous and

$$\|\nabla_x F(y, s) - \nabla_x F(x, s)\| \leq \sigma \|y - x\|^\alpha, \quad (4)$$

for all $x, y \in B(x_0, \delta)$ and all $s \in I(x) \cup I(y)$, where

$$I(x) = \{s^* \in S : f(x) = F(x, s^*)\}. \quad (5)$$

We say that f is lower- $C^{1,\alpha}$ on U (and we denote $f \in LC^{1,\alpha}$) if the above definition is fulfilled at every $x \in U$. Removing condition (4) from Definition 2 or setting $\alpha = 0$, we obtain the definition of the lower- C^1 function given in the introduction. Hence, the above definition is a strengthening of the lower- C^1 property. In Subsection 3.3 we provide an example of a LC^1 function that is not $LC^{1,\alpha}$ for any $\alpha > 0$ (see Proposition 13).

Similarly to Definition 2, the following notion strengthens the notion of approximate convexity defined in (1).

Definition 3 (α -weakly convex function) Let U be a nonempty open subset of \mathbb{R}^n and $0 < \alpha \leq 1$. A locally Lipschitz function $f : U \rightarrow \mathbb{R}$ is called α -weakly convex at $x_0 \in U$, if there exist $\sigma > 0$ and $\delta > 0$ such that for all $x, y \in B(x_0, \delta)$ and $t \in (0, 1)$

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) + \sigma t(1-t)\|x - y\|^{1+\alpha}. \quad (6)$$

The function f is called α -weakly convex, if it is α -weakly convex at every $x \in U$.

Remark 4 Taking $\alpha = 1$ in the above definition corresponds to the notion of weak convexity, see (2). On the other hand, the value $\alpha = 0$ has no practical interest. It yields a notion which is strictly weaker than approximate convexity (since “for every $\varepsilon > 0$ ” has been replaced by “there exists $\sigma > 0$ ”) and which does not ensure the Clarke regularity of the function.

Finally we need the notion of α -hypomonotone operator, which lies strictly between submonotonicity and hypomonotonicity.

Definition 5 (α -hypomonotone operator) Let U be a nonempty open subset of \mathbb{R}^n and $0 < \alpha \leq 1$. A multivalued mapping $T : U \rightrightarrows \mathbb{R}^n$ is called α -hypomonotone at $x_0 \in U$, if there exist $\sigma > 0$ and $\delta > 0$ such that for all $x, y \in B(x_0, \delta)$, $x^* \in \partial f(x)$ and $y^* \in \partial f(y)$ we have

$$\langle y^* - x^*, y - x \rangle \geq -\sigma \|y - x\|^{1+\alpha}. \quad (7)$$

The operator T is called α -hypomonotone, if it is α -hypomonotone at every $x \in U$.

Remark 6 An analogous remark applies here. Setting $\alpha = 1$ we recover the notion of hypomonotonicity, while the value $\alpha = 0$ has no interest for our purposes.

3 Main results

In Subsection 3.1 we establish subdifferential and mixed characterizations of the class of lower- $C^{1,\alpha}$ functions, while in Subsection 3.2 we show the coincidence of that class with the class of locally Lipschitz α -weakly convex functions and give an epigraphical characterization. These results are in the spirit of [22], [5], [15] (for approximately convex functions) and of [18], [4], [2] (for weakly convex functions). We also quote [4] and [1] for a study of epigraphical properties of such functions.

In Subsection 3.3 we give a complete classification of the aforementioned classes and examples distinguishing them. We also present subclasses with a particular interest in optimization.

3.1 Subdifferential characterizations

The following result is an expected characterization of α -weak convexity.

Theorem 7 (characterizations) *Let U be an open set of \mathbb{R}^n and $f : U \rightarrow \mathbb{R}$ a locally Lipschitz function. The following statements are equivalent:*

- (i) f is α -weakly convex on U ;
- (ii) ∂f is α -hypomonotone on U ;

(iii) for all $x_0 \in U$, there exist $\sigma, \delta > 0$ such that for all $x \in B(x_0, \delta)$, $x^* \in \partial f(x)$, and $u \in \mathbb{R}^n$ with $x + u \in B(x_0, \delta)$,

$$f(x + u) \geq f(x) + \langle x^*, u \rangle - \sigma \|u\|^{1+\alpha}. \quad (8)$$

Proof. (i) \Rightarrow (iii). Fix $x_0 \in U$, $\sigma > 0$, $\delta > 0$ given by Definition 3. Let us consider any $x \in B(x_0, \delta)$ and $u \in \mathbb{R}^n$ such that $x + u \in B(x_0, \delta)$. Then for $z \in B(x_0, \delta)$ sufficiently closed to x and such that $z + u \in B(x_0, \delta)$, one has

$$f(z + tu) \leq tf(z + u) + (1 - t)f(z) + \sigma t(1 - t) \|u\|^{1+\alpha}$$

or equivalently

$$\frac{f(z + tu) - f(z)}{t} \leq f(z + u) - f(z) + \sigma(1 - t) \|u\|^{1+\alpha}$$

Taking the “limsup” when $z \rightarrow x$ and $t \rightarrow 0+$ in both sides, one gets

$$f^o(x; u) \leq f(x + u) - f(x) + \sigma \|u\|^{1+\alpha}$$

which in view of (3) yields the result.

(iii) \Rightarrow (ii). Fix $x_0 \in U$, $\sigma > 0$, $\delta > 0$ and take any $x, y \in B(x_0, \delta)$, $x^* \in \partial f(x)$ and $y^* \in \partial f(y)$. Then one has

$$f(y) \geq f(x) + \langle x^*, y - x \rangle - \sigma \|x - y\|^{1+\alpha} \quad \text{and} \quad f(x) \geq f(y) + \langle y^*, x - y \rangle - \sigma \|x - y\|^{1+\alpha}$$

which by addition yields

$$\langle x^* - y^*, x - y \rangle \geq -2\sigma \|x - y\|^{1+\alpha}.$$

This shows the α -hypomonotonicity of ∂f .

(ii) \Rightarrow (i). Suppose ∂f is α -hypomonotone and let $\sigma > 0$, $\delta > 0$ as in Definition 5. Fix $x_1, x_2 \in B(x_0, \delta)$ and for any $t \in (0, 1)$ set $x_t = tx_1 + (1 - t)x_2$ so that

$$x_t - x_1 = (1 - t)(x_2 - x_1) \quad \text{and} \quad x_t - x_2 = t(x_1 - x_2). \quad (9)$$

By the Lebourg mean value theorem (see [12] or [3, Theorem 2.3.7]), for every $i \in \{1, 2\}$ there exists $z_i \in [x_i, x_t[$ and $z_i^* \in \partial f(z_i)$ such that

$$f(x_t) = f(x_i) + \langle z_i^*, x_t - x_i \rangle. \quad (10)$$

Multiplying (10) respectively by t for $i = 1$ and by $(1 - t)$ for $i = 2$ and adding the resulting inequalities we conclude in view of (9) that

$$f(x_t) = tf(x_1) + (1 - t)f(x_2) - t(1 - t)\langle z_1^* - z_2^*, x_1 - x_2 \rangle. \quad (11)$$

Since

$$\frac{x_1 - x_2}{\|x_1 - x_2\|} = \frac{z_1 - z_2}{\|z_1 - z_2\|},$$

the definition of α -hypomonotonicity implies

$$\langle z_1^* - z_2^*, x_1 - x_2 \rangle \geq -\sigma \|z_1 - z_2\|^\alpha \|x_1 - x_2\| \geq -\sigma \|x_1 - x_2\|^{1+\alpha},$$

so (11) yields

$$f(x_t) \leq tf(x) + (1 - t)f(y) + \sigma t(1 - t) \|x - y\|^{1+\alpha},$$

which ends the proof. \square

Let us note that the property that f is locally Lipschitz is only used for the implication (ii) \Rightarrow (i), in which the Lebourg mean value theorem for locally Lipschitz functions was needed. All other implications can be adapted to the case that f is lower semicontinuous and ∂f is its Clarke-Rockafellar subdifferential (we refer to [3] or [4] for the corresponding definition).

3.2 Coincidence of α -weakly convex and $LC^{1,\alpha}$ functions

Let us now show the coincidence of the classes of locally Lipschitz α -weakly convex functions (Definition 3) and of $LC^{1,\alpha}$ functions (Definition 2). This result comes to complete statements of similar nature, previously established in [5, Corollary 3] (for approximately convex functions) and in [18], [23] (for weakly convex functions).

Theorem 8 (coincidence result) *Let U be a nonempty open subset of \mathbb{R}^n and let $0 < \alpha \leq 1$. Then a locally Lipschitz function $f : U \rightarrow \mathbb{R}$ is lower- $C^{1,\alpha}$ if and only if f is α -weakly convex.*

Proof. (\Rightarrow). Let us assume that f is lower- $C^{1,\alpha}$ and let us fix any $x_0 \in U$. Then let us consider $\delta, \sigma > 0$, a nonempty compact set S and a continuous function $F(x, s)$ according to the Definition 2 so that

$$f(x) = \max_{s \in S} F(x, s), \quad \text{for all } x \in B(x_0, \delta),$$

and

$$\|\nabla F(y, s) - \nabla F(x, s)\| \leq \sigma \|y - x\|^\alpha, \quad (12)$$

for all $x, y \in B(x_0, \delta)$ and $s \in I(x) \cup I(y)$. Let $x \in B(x_0, \delta)$ and $u \in \mathbb{R}^n$ be such that $x + u \in B(x_0, \delta)$ and set $y = x + u$. Since S and $I(x)$ are compact, it follows (see [19, Theorem 10.31]) that

$$\partial f(x) = \text{co} \{ \nabla F(x, s), s \in I(x) \},$$

where $\text{co}(A)$ denotes the convex hull of a set A . For any $x^* \in \partial f(x)$, by the Caratheodory theorem, there exist $\lambda_1, \dots, \lambda_{n+1}$ in \mathbb{R}_+ with $\sum_i \lambda_i = 1$ and s_1, \dots, s_{n+1} in $I(x)$ such that

$$x^* = \sum_{i=1}^{n+1} \lambda_i \nabla F(x, s_i).$$

Applying for every $i \in \{1, \dots, n+1\}$ the classical mean-value theorem to the differentiable function $x \mapsto F(x, s_i)$ we obtain $z_i \in [x, y[$ such that

$$F(y, s_i) - F(x, s_i) = \langle \nabla F(z_i, s_i), y - x \rangle.$$

Since $s_i \in I(x)$, we have successively

$$\begin{aligned} f(y) &\geq F(y, s_i) \\ &= F(x, s_i) + \langle \nabla F(z_i, s_i), y - x \rangle \\ &= f(x) + \langle \nabla F(x, s_i), y - x \rangle + \langle \nabla F(z_i, s_i) - \nabla F(x, s_i), y - x \rangle. \end{aligned}$$

Multiplying by $\lambda_i \geq 0$ and adding the resulting inequalities for $i \in \{1, \dots, n+1\}$ we obtain (recalling $y = x + u$) that

$$f(x + u) \geq f(x) + \langle x^*, u \rangle + \sum_{i=1}^{n+1} \lambda_i \langle \nabla F(z_i, s_i) - \nabla F(x, s_i), u \rangle. \quad (13)$$

Since $s_i \in I(x)$ for $i \in \{1, \dots, n+1\}$, relation (12) yields

$$\langle \nabla F(x, s_i) - \nabla F(z_i, s_i), u \rangle \leq \sigma \|u\| \|z_i - x\|^\alpha.$$

Since $z_i \in [x, y[$ this yields

$$\langle \nabla F(x, s_i) - \nabla F(z_i, s_i), u \rangle \leq \sigma \|u\| \|y - x\|^\alpha = \sigma \|u\|^{1+\alpha}.$$

Replacing into (13) we get

$$f(x + u) \geq f(x) + \langle x^*, u \rangle - \sigma \|u\|^{1+\alpha},$$

so the assertion follows from Theorem 7 (iii) \Rightarrow (i).

(\Leftarrow). Conversely, let us assume f is α -weakly convex and let us consider $x_0 \in U$. Then for some $\sigma, \delta > 0$ and all $y, z \in B(x_0, \delta)$, $z^* \in \partial f(z)$ we have

$$f(y) \geq f(z) + \langle z^*, y - z \rangle - \sigma \|y - z\|^{1+\alpha}. \quad (14)$$

Taking eventually $\tilde{\sigma} > \sigma$, we may assume that the above inequality is strict for all $y \neq z \in B(x_0, \delta)$ and all $z^* \in \partial f(z)$. Set

$$S = \left\{ (z, z^*) \in \mathbb{R}^n \times \mathbb{R}^n, \quad \|z - x_0\| \leq \frac{\delta}{2}, \quad z^* \in \partial f(z) \right\}$$

Since ∂f is locally bounded and has a closed graph (see [3, Proposition 2.1.5], for example) it follows that S is compact. Moreover, S is nonempty since it contains the set $\{x_0\} \times \partial f(x_0)$. Let us now define

$$\begin{aligned} F : B(x_0, \delta/2) \times S &\longrightarrow \mathbb{R} \\ (x, (z, z^*)) &\longmapsto F(x, (z, z^*)) := f(z) + \langle z^*, x - z \rangle - \sigma \|x - z\|^{1+\alpha}. \end{aligned}$$

Then for every $x \in B(x_0, \delta/2)$ and every $s = (z, z^*) \in S$ we have in view of (14) (and the choice of $\sigma > 0$) that

$$f(x) \geq F(x, (z, z^*))$$

with strict inequality whenever $x \neq z$. Thus for every $x \in B(x_0, \delta/2)$

$$f(x) = \max_{(z, z^*) \in S} F(x, (z, z^*)),$$

and

$$I(x) = \{x\} \times \partial f(x).$$

Note also that

$$\nabla_x F(x, (z, z^*)) = \begin{cases} z^* - \sigma(1 + \alpha) \|x - z\|^{\alpha-1} (x - z) & \text{if } x \neq z \\ z^* & \text{if } x = z \end{cases}$$

Let now any $x, y \in B(x_0, \delta)$ and $s = (z, z^*) \in I(x) \cup I(y)$. It follows that $z \in \{x, y\}$. Let us suppose (with no loss of generality) that $z = y$. Then

$$\|\nabla_x F(y, s) - \nabla_x F(x, s)\| = \sigma(1 + \alpha) \|y - x\|^\alpha.$$

Thus (4) of Definition 2 holds. To complete the proof, it suffices to check the continuity of $\nabla_x F(x, (z, z^*))$ on $B(x_0, \delta) \times S$. This is clear at every point $(x, (z, z^*))$ with $x \neq z$, so let us suppose that $x = z$, that is, $(x, (z, z^*)) = (x, (x, z^*))$ and let $(x_n, (z_n, z_n^*))_{n \geq 1}$ be a sequence of $B(x_0, \delta) \times S$ converging to $(x, (x, z^*))$. For all $n \in \mathbb{N}$ such that $x_n \neq z_n$ we have

$$\begin{aligned} &\|\nabla_x F(x, (x, z^*)) - \nabla_x F(x_n, (z_n, z_n^*))\| \\ &= \left\| z^* - z_n^* + \sigma(1 + \alpha) \|x_n - z_n\|^{\alpha-1} (x_n - z_n) \right\| \\ &\leq \|z^* - z_n^*\| + \sigma(1 + \alpha) \|x_n - z_n\|^\alpha. \end{aligned}$$

On the other hand, for all $n \in \mathbb{N}$ such that $x_n = z_n$ we have

$$\|\nabla_x F(x, (x^*, z^*)) - \nabla_x F(x_n, (x_n, z_n^*))\| = \|z^* - z_n^*\|.$$

Thus, it follows easily that

$$\|\nabla_x F(x, (x, z^*)) - \nabla_x F(x_n, (z_n, z_n^*))\| \longrightarrow 0$$

as $(x_n, (z_n, z_n^*)) \longrightarrow (x, (x, z^*))$. This shows that $\nabla_x F$ is jointly continuous, so $f \in LC^{1, \alpha}$. \square

Remark 9 ($LC^{1,1} \equiv LC^2$) Taking $\alpha = 1$ in the above proof we obtain that the class of the lower- $C^{1,1}$ functions and of the locally Lipschitz weakly convex functions coincide. In view of the classical result of Rockafellar [18] (recalled in the introduction), we conclude that the classes $LC^{1,1}$ and LC^2 coincide.

Let us now provide a characterization of the epigraphs of $LC^{1,\alpha}$ functions, in terms of the truncated normal cone operator. We first recall the definition of the latter: if C is a nonempty subset of a Euclidean space \mathbb{R}^m ($m \in \mathbb{N}^*$), then the (Clarke) normal cone of C at $u \in C$ is defined by

$$N_C(u) = \{u^* \in \mathbb{R}^m : \langle u^*, v \rangle \leq 0, \forall v \in T_C(u)\}, \quad (15)$$

where the Clarke tangent cone $T_C(u)$ is defined as follows:

$$v \in T_C(u) \iff \begin{cases} \forall \varepsilon > 0, \exists \delta > 0 \text{ such that} \\ \forall u' \in B(u, \delta) \cap C, \forall t \in]0, \delta[, (u' + tB(v, \varepsilon)) \cap C \neq \emptyset. \end{cases} \quad (16)$$

We put $N_C(u) = \emptyset$, whenever $u \notin C$. For any $r > 0$ we denote by $N_C^r(u)$ the truncated Clarke normal cone, that is,

$$N_C^r(u) = N_C(u) \cap B[0, r],$$

where $B[0, r]$ denotes the closed ball in \mathbb{R}^m of center 0 and radius r . We further denote by

$$\text{epi } f := \{(x, \beta) \in \mathbb{R}^{n+1} : \beta \geq f(x)\}$$

the epigraph of the function f defined on \mathbb{R}^n . By [3, p. 56], for all $u_0 = (x_0, f(x_0)) \in \text{epi } f$ we have

$$N_{\text{epi } f}(u_0) = \mathbb{R}^+ (\partial f(x_0), -1).$$

Let us finally note that, if f is κ -Lipschitz on a ball B of \mathbb{R}^n , then for all x_1, x_2 in B , we have

$$\|x_2 - x_1\| \leq \|u_2 - u_1\| \leq \sqrt{1 + \kappa^2} \|x_2 - x_1\|, \quad (17)$$

where $u_i := (x_i, f(x_i))$, $i \in \{1, 2\}$ and where we use the same notation to denote the Euclidean norm of the spaces \mathbb{R}^n and \mathbb{R}^{n+1} .

The following result is analogous to the ones established in [4, Section 5] (for LC^2 functions) and in [1, Theorem 4.1.4] (for LC^1 functions).

Corollary 10 (epigraphical characterization) *Let $f : U \rightarrow \mathbb{R}$ be a locally Lipschitz function defined on an open subset U of \mathbb{R}^n . The following two assertions are equivalent:*

- (i) *the function f is lower- $C^{1,\alpha}$;*
- (ii) *the operator $N_{\text{epi } f}^1 : \mathbb{R}^{n+1} \rightrightarrows \mathbb{R}^{n+1}$ is α -hypomonotone.*

Proof. (i) \Rightarrow (ii) Let $u_0 \in \text{epi } f$. We can suppose without loss of generality that $u_0 = (x_0, f(x_0))$ for $x_0 \in U$ (otherwise $N_{\text{epi } f}(u)$ is reduced to $\{0\}$ for all u in a neighborhood of u_0 , so that (7) is clearly satisfied).

Let now $\kappa, \delta_1 > 0$ such that f is κ -Lipschitz on $B(x_0, \delta_1)$. By Theorem 8, the function f is weakly convex, so Theorem 7 (i) \Rightarrow (iii) yields that there exist $\delta_2 > 0$ and $\sigma > 0$ such that for all $x_1, x_2 \in B(x_0, \delta_2)$, $x_1^* \in \partial f(x_1)$ and $x_2^* \in \partial f(x_2)$

$$f(x_2) - f(x_1) \geq \langle x_1^*, x_2 - x_1 \rangle - \sigma \|x_1 - x_2\|^{1+\alpha}. \quad (18)$$

Set $\delta = \min\{\delta_1, \delta_2\}$ and take $u_1, u_2 \in B(u_0, \delta) \cap \text{epi } f$ (we use the same notation $B(u_0, \delta)$ to denote the ball of center u_0 and radius $\delta > 0$ in the space \mathbb{R}^{n+1}). In particular, u_1 has the form (x_1, β_1) with $\beta_1 \geq f(x_1)$. There are two cases:

- If $\beta_1 > f(x_1)$, then $N_{\text{epi } f}^1(u_1) = \{0\}$.

- If $\beta_1 = f(x_1)$, then

$$N_{\text{epi } f}^1(u_1) = \mathbb{R}^+(\partial f(x_1), -1) \cap B[0, 1].$$

So for every $u_1^* \in N_{\text{epi } f}^1(u_1)$, there exists $x_1^* \in \partial f(x_1)$ such that $u_1^* = \mu_1(x_1^*, -1)$. Note also that we can bound μ_1 uniformly. Since f is κ -Lipschitz on $B(x_0, \delta)$, one has $\|x_1^*\| \leq \kappa$ (see [3, Proposition 2.1.2], for example). As $\|u_1^*\| \leq 1$, one obtains $\mu_1 \leq (1 + \kappa^2)^{-\frac{1}{2}}$.

Since $\beta_2 \geq f(x_2)$, (18) implies

$$\langle (x_1^*, -1), (x_2 - x_1, \beta_2 - \beta_1) \rangle \leq \sigma \|x_1 - x_2\|^{1+\alpha}.$$

Here again we use the same notation for the scalar products in \mathbb{R}^n and in \mathbb{R}^{n+1} . In particular, $\langle (x, \alpha), (y, \beta) \rangle := \langle x, y \rangle + \alpha\beta$, for all $x, y \in \mathbb{R}^n$ and all $\alpha, \beta \in \mathbb{R}$.

In both cases, for every $u_1^* \in N_{\text{epi } f}^1(u_1)$ we have

$$\langle u_1^*, u_2 - u_1 \rangle \leq (1 + \kappa^2)^{-\frac{1}{2}} \sigma \|x_1 - x_2\|^{1+\alpha},$$

which in view of (17) yields

$$\langle u_1^*, u_2 - u_1 \rangle \leq (1 + \kappa^2)^{-\frac{1}{2}} \sigma \|u_1 - u_2\|^{1+\alpha}.$$

Interchanging the roles of u_1 and u_2 , for every $u_2^* \in N_{\text{epi } f}^1(u_2)$ we have

$$\langle u_2^*, u_2 - u_1 \rangle \geq -(1 + \kappa^2)^{-\frac{1}{2}} \sigma \|u_1 - u_2\|^{1+\alpha}.$$

Subtracting the last two equations, we get

$$\langle u_2^* - u_1^*, u_2 - u_1 \rangle \geq -2(1 + \kappa^2)^{-\frac{1}{2}} \sigma \|u_1 - u_2\|^{1+\alpha},$$

which means that $N_{\text{epi } f}^1$ is α -hypomonotone.

(ii) \Rightarrow (i) Fix $x_0 \in U$ and set $u_0 = (x_0, f(x_0))$. Let δ_1 and σ such that for all $u_1, u_2 \in B(x_0, \delta_1)$, $u_1^* \in N_{\text{epi } f}^1(u_1)$ and $u_2^* \in N_{\text{epi } f}^1(u_2)$

$$\langle u_2^* - u_1^*, u_2 - u_1 \rangle \geq -\sigma \|u_1 - u_2\|^{1+\alpha}. \quad (19)$$

Let δ_2 and κ be such that f is κ -Lipschitz on $B(x_0, \delta_1)$ and set

$$\delta = \frac{\min\{\delta_1, \delta_2\}}{\sqrt{1 + \kappa^2}}.$$

Let $x_1, x_2 \in B(x_0, \delta)$, $x_1^* \in \partial f(x_1)$ and $x_2^* \in \partial f(x_2)$. For $i \in \{1, 2\}$, set $u_i = (x_i, f(x_i))$ and $u_i^* := (1 + \kappa^2)^{-\frac{1}{2}}(x_i^*, -1)$. Observe that $u_i \in B(u_0, \delta_1)$ and $u_i^* \in N_{\text{epi } f}^1(u_i)$. Thus (19) can be rephrased as

$$\langle (x_2^* - x_1^*, 0), (x_2 - x_1, f(x_2) - f(x_1)) \rangle \geq -\sigma(1 + \kappa^2)^{\frac{1}{2}} \|u_1 - u_2\|^{1+\alpha}.$$

Using (17) we get

$$\langle x_2^* - x_1^*, x_2 - x_1 \rangle \geq -\sigma(1 + \kappa^2) \|x_1 - x_2\|^{1+\alpha}.$$

Thus ∂f is α -hypomonotone. By Theorem 7 (ii) \Rightarrow (i) and Theorem 8, we conclude that f is $LC^{1,\alpha}$. \square

3.3 Classification

Let us fix a nonempty open subset U of \mathbb{R}^n and let us consider the following two particular classes of functions.

– **(locally decomposable functions)** We say that a locally Lipschitz function $f : U \rightarrow \mathbb{R}$ is *locally decomposable* on U as a sum of a convex function and a $C^{1,\alpha}$ function if for all $x_0 \in U$ there exists $\delta > 0$, a convex continuous function $k : B(x_0, \delta) \rightarrow \mathbb{R}$ and a $C^{1,\alpha}$ -function $h : B(x_0, \delta) \rightarrow \mathbb{R}$ (that is, h is differentiable with α -Hölder derivative) such that

$$f(x) = k(x) + h(x), \quad \text{for all } x \in B(x_0, \delta).$$

– **(locally composite functions)** We say that a locally Lipschitz function $f : U \rightarrow \mathbb{R}$ is *locally composite* on U , if for every $x_0 \in U$ there exists $\delta > 0$, a convex continuous function $g : \mathbb{R}^n \rightarrow \mathbb{R}$ and a $C^{1,\alpha}$ -function $G : B(x_0, \delta) \rightarrow \mathbb{R}^n$ such that

$$f(x) = g(G(x)), \quad \text{for all } x \in B(x_0, \delta).$$

This implies (see [19, p. 445], for example) that

$$\partial f(x) = \nabla G(x)^* \partial g(G(x)), \quad \text{for all } x \in B(x_0, \delta).$$

Proposition 11 *Let $f : U \rightarrow \mathbb{R}$ be a locally Lipschitz function and $0 < \alpha < 1$. Consider the following conditions:*

- (i) *f is locally decomposable on U as a sum of a convex continuous and a $C^{1,\alpha}$ function ;*
- (ii) *f is locally composite on U with a convex continuous and a $C^{1,\alpha}$ function ;*
- (iii) *f is a $LC^{1,\alpha}$ function.*

Then (i) \implies (ii) \implies (iii).

Proof. (i) \implies (ii). Having a local decomposition $f = k + h$, set $g(x, r) = k(x) + r$ for $(x, r) \in \mathbb{R}^n \times \mathbb{R}$ and $G(x) = (x, h(x))$ for $x \in \mathbb{R}^n$. It is straightforward to see that $f(x) = g(G(x))$, that G is $C^{1,\alpha}$ and that g is convex and continuous.

(ii) \implies (iii). Let $x_0 \in U$, $\delta > 0$ and $g, h : B(x_0, \delta) \rightarrow \mathbb{R}$, g being convex continuous and $G \in C^{1,\alpha}(B(x_0, \delta))$ such that $f(x) = g(G(x))$ for all $x \in B(x_0, \delta)$. For all x near x_0 , one has

$$\partial f(x) = \nabla G(x)^* \partial g(G(x))$$

Since ∇G is α -Hölderian, let $\sigma > 0$ such that for all $x, y \in B(x_0, \delta)$

$$\|\nabla G(y) - \nabla G(x)\| \leq \sigma \|y - x\|^\alpha. \quad (20)$$

Let $x, y \in B(x_0, \delta)$. For any $x^* \in \partial f(x)$, there exists $\zeta \in \partial g(G(x))$ such that $x^* = \nabla G(x)^* \zeta$. Since g is convex, it follows that

$$f(y) - f(x) = g(G(y)) - g(G(x)) \geq \langle \zeta, G(y) - G(x) \rangle. \quad (21)$$

Applying the mean value theorem to the function G on the segment $[x, y]$ we obtain $z \in [x, y]$ such that

$$G(y) - G(x) = \nabla G(z)(y - x). \quad (22)$$

By (20), it holds

$$\|\nabla G(z) - \nabla G(x)\| \leq \sigma \|z - x\|^\alpha \leq \sigma \|y - x\|^\alpha. \quad (23)$$

Thus by (21), (22) and (23), we can write

$$\begin{aligned} f(y) - f(x) &\geq \langle \zeta, \nabla G(z)(y - x) \rangle \\ &= \langle \zeta, \nabla G(x)(y - x) \rangle + \langle \zeta, (\nabla G(z) - \nabla G(x))(y - x) \rangle \\ &\geq \langle \zeta, \nabla G(x)(y - x) \rangle - \sigma \|\zeta\| \|y - x\|^{1+\alpha} \end{aligned}$$

Moreover, there exists a constant $\kappa > 0$ which bounds uniformly the norm of every subgradient of the convex continuous function g near x_0 . Thus it holds

$$f(y) - f(x) \geq \langle x^*, y - x \rangle - \sigma \kappa \|y - x\|^{1+\alpha},$$

and we can conclude by Theorem 7(iii) \Rightarrow (i) and Theorem 8. \square

Remark 12 (conjecture) A classical result of Rockafellar [18] (see also [23], [8]) asserts that every LC^2 function is decomposable as a sum of a convex continuous and a concave quadratic function. Moreover, in view of Remark 9, the classes $LC^{1,1}$ and LC^2 coincide. Thus, in case $\alpha = 1$, the three assertions of Proposition 11 are then equivalent. It is not known if an analogous equivalence holds for the classes of LC^1 and $LC^{1,\alpha}$ functions.

Let us now give an example of a LC^1 function f , which does not belong to any of the classes $LC^{1,\alpha}$ for $\alpha > 0$. More precisely, we have the following proposition.

Proposition 13

$$\bigcup_{0 < \alpha < 1} LC^{1,\alpha} \subsetneq LC^1$$

Proof. The inclusion follows directly from Definition 2. To see that the inclusion is strict, let us consider the function $f : \mathbb{R} \rightarrow \mathbb{R}$ defined as follows:

$$f(x) = - \int_0^x g(t) dt,$$

where

$$g(t) = \begin{cases} 0 & t \leq 0 \\ \frac{1}{|\ln t|} & t > 0. \end{cases}$$

It is easily seen that g is continuous on \mathbb{R} , so that f is of class C^1 . In particular, $f \in LC^1$. Note also that $f(0) = 0$ and $f'(0) = 0$.

Let us prove that for any $\alpha > 0$ the function f does not belong to the class $LC^{1,\alpha}$. Indeed, suppose towards a contradiction that there exists $\alpha > 0$ such that $f \in LC^{1,\alpha}$. Then by Theorem 8 and Theorem 7 (i) \Rightarrow (iii) there exist $\sigma, \delta > 0$ such that for all $x \in (0, 1)$,

$$f(x) \geq -\sigma |x|^{1+\alpha}.$$

Set now $\phi(x) = f(x) + \sigma |x|^{1+\alpha}$. Then the function ϕ is C^1 , non-negative and $\phi(0) = 0$. It follows easily that there exists a sequence $(x_n)_{n \geq 1}$ of positive real numbers converging to 0 such that $\phi'(x_n) \geq 0$. (Indeed, if for some $\delta > 0$ we have $\phi'(x) < 0$ for all $x \in (0, \delta)$, then ϕ should necessarily take negative values.) We compute $\phi'(x) = (1 + \alpha)\sigma x^\alpha - g(x)$ for $x > 0$. Then we have for all $n > 0$

$$(1 + \alpha)\sigma \geq \frac{1}{x_n^\alpha |\ln x_n|}.$$

Since $\alpha > 0$ the right-hand side tends to $+\infty$ when n grows. We thus obtain a contradiction. It follows that

$$f \in C^1 \setminus \bigcup_{\alpha > 0} LC^{1,\alpha},$$

which proves the assertion. \square

Let us complete our classification with the following proposition.

Proposition 14

$$LC^2 \subsetneq \bigcap_{0 < \alpha < 1} LC^{1,\alpha}$$

Proof. Since every LC^2 function is a fortiori $LC^{1,\alpha}$ for all $0 < \alpha < 1$, the inclusion holds. To see that the inclusion is strict, let us consider the function

$$f(x) = \int_0^x g(t) dt, \quad \text{for all } x \in \mathbb{R},$$

where

$$g(t) = \begin{cases} 0 & t \leq 0 \\ t \ln t & t > 0. \end{cases}$$

Then g is continuous on \mathbb{R} and clearly not Lipschitz around $t = 0$. Let us show that, for any $0 < \alpha < 1$, g is α -Hölderian in a neighborhood of 0. To this end, take x, y sufficiently small to ensure that are inside a neighborhood of 0 in which g is decreasing. We can suppose without loss of generality that $y < x$. We may suppose $x > 0$ (else the condition of α -Hölderianity is trivially fulfilled), and we distinguish three cases.

Case 1. $y \leq 0$. Then we can write

$$\frac{|g(x) - g(y)|}{|x - y|^\alpha} = \frac{x |\ln x|}{|x - y|^\alpha} \leq \frac{x |\ln x|}{x^\alpha} = x^{1-\alpha} |\ln x|. \quad (24)$$

Case 2. $0 < y < x/2$. In this case $0 > g(y) > g(x)$ so that

$$\frac{|g(x) - g(y)|}{|x - y|^\alpha} \leq \frac{|g(x)|}{|x/2|^\alpha} \leq 2^\alpha |\ln x| x^{1-\alpha}. \quad (25)$$

Case 3. $x/2 < y < x$. Applying the mean-value theorem for the function g to the segment $[x, y]$ (where g is C^∞) we obtain $z \in [x, y]$ such that

$$\frac{|g(x) - g(y)|}{|x - y|^\alpha} \leq (|\ln z| + 1) |x - y|^{1-\alpha} \leq (|\ln \frac{x}{2}| + 1) x^{1-\alpha}. \quad (26)$$

In all cases (24)-(26), the quantity $|x - y|^{-\alpha} |g(x) - g(y)|$ is bounded when x and y are sufficiently close to 0. Thus, there exist $\delta > 0$ and $M > 0$ such that for all $x, y \in]-\delta, \delta[$ with $x \neq y$ we have

$$\frac{|g(x) - g(y)|}{|x - y|^\alpha} \leq M.$$

This means that g is α -Hölderian on $] -\delta, \delta[$.

It follows that f is C^1 on \mathbb{R} and locally $C^{1,\alpha}$ around 0, for any $0 < \alpha < 1$. We us prove that f is not LC^2 around 0. To this end, let us assume, towards a contradiction, that there exists $\delta > 0$ such that ∂f is hypomonotone on $B(x_0, \delta)$. Since f is C^1 , we have $\partial f(x) = \{g(x)\}$ for all $x \in \mathbb{R}$, and in particular $\partial f(0) = \{0\}$. Then for all $\sigma > 0$ and $x \in B(x_0, \delta)$,

$$x g(x) \geq -\sigma |x|^2.$$

This implies

$$\ln x \geq -\sigma \quad \text{for all } 0 < x < \delta,$$

which is a clear contradiction. □

Let us resume the results in the following diagram.

$$\begin{array}{c}
LC^\infty = \underset{(2 < k < +\infty)}{LC^k} = LC^2 = LC^{1,1} \subsetneq \underset{(0 < \alpha < 1)}{LC^{1,\alpha}} \subsetneq LC^1 \\
\\
\bigcup_{0 < \alpha < 1} LC^{1,\alpha} \subsetneq LC^1 \\
\\
LC^2 \subsetneq \bigcap_{0 < \alpha < 1} LC^{1,\alpha}
\end{array}$$

Acknowledgment: The first author is indebted to L. Thibault for useful discussions.

References

- [1] AUSSEL, D., DANILIDIS, A. & THIBAUT, L., Subsmooth sets: functional characterizations and related concepts, *Trans. Amer. Math. Soc.* (to appear).
- [2] BERNARD, F. & THIBAUT, L., Uniform Prox-regularity of functions and epigraphs in Hilbert spaces, preprint 25p, Montpellier, 2002.
- [3] CLARKE, F.H., Optimization and Nonsmooth Analysis (Wiley Interscience, New York 1983).
- [4] CLARKE, F.H., STERN, R. & WOLENSKI, P., Proximal Smoothness and the Lower- C^2 Property, *J. Convex Anal.* **2** (1995), 117-144.
- [5] DANILIDIS, A. & GEORGIEV, P., Approximate convexity and submonotonicity, *J. Math. Anal. Appl.* **291** (2004), 292-301.
- [6] GEORGIEV, P., Submonotone Mappings in Banach Spaces and Applications, *Set-Valued Analysis* **5** (1997), 1-35.
- [7] GREEN, J.-W., Approximately convex functions, *Duke Math. J.* **19** (1952), 499-504.
- [8] HIRIART-URRUTY, J.-B., "Generalized differentiability, duality and optimization for problems dealing with differences of convex functions", in: Convexity and Duality in Optimization", *Lecture Notes in Econom. Math. Systems* **256** (1984), 37-70.
- [9] HYERS, D.-H. & ULAM, S.-M., Approximately convex functions, *Proc. Amer. Math. Soc.* **3** (1952), 821-828.
- [10] JANIN, R., Sur des multiapplications qui sont des gradients généralisés, *C.R. Acad. Sc. Paris* **294** (1982), 117-119.
- [11] JOURANI, A., Subdifferentiability and subdifferential monotonicity of γ -paraconvex functions, *Control Cybernet.* **25** (1996), 721-737.
- [12] LEBOURG, G., Generic differentiability of Lipschitzian functions, *Trans. Amer. Math. Soc.* **256** (1979), 125-144.

- [13] MALIVERT, C., PENOT, J.-P. & THERA, M., Minimisation d'une fonction régulière sur un fermé non-régulier et non-convexe d'un espace de Hilbert, *C. R. Acad. Sc. Paris* **286** (Série A) (1978), 1191-1193.
- [14] NGAI, H.V., LUC, D.T. & THERA, M., "On ε -convexity and ε -monotonicity", in: Calculus of Variations and Differential Equations, A. Ioffe, S. Reich, and I. Shafrir (Eds), Research Notes in Mathematical Series (Chapman & Hall, 82-100, 1999).
- [15] NGAI, H.V., LUC, D.T. & THERA, M., Approximate convex functions, *Journal of Nonlinear and Convex Analysis* **1** (2000), 155-176.
- [16] PALES, Z., On approximately convex functions, *Proc. Amer. Math. Soc.* **131** (2003), 243-252.
- [17] PENOT, J.-P., Favorable Classes of Mappings and Multimappings in Nonlinear Analysis and Optimization, *J. Convex Anal.* **3** (1996) 97-116.
- [18] ROCKAFELLAR, R.T., "Favorable Classes of Lipschitz Continuous Functions in Subgradient Optimization" in: *Non-differentiable Optimization* (1982), Nurminski E. (eds), Pergamon Press, New York.
- [19] ROCKAFELLAR, R.T. & WETS, R., *Variational Analysis*, Grundlehren der Mathematischen, Wissenschaften, Vol. **317**, (Springer, 1998).
- [20] ROLEWICZ, S., On the coincidence of some subdifferentials in the class of $\alpha(\cdot)$ -paraconvex functions, *Optimization* **50** (2001), 353-363.
- [21] POLIQUIN, R. & ROCKAFELLAR, T., Prox-regular functions in variational analysis, *Trans. Amer. Math. Soc.* **348** (1996), 1805-1838.
- [22] SPINGARN, J.E., Submonotone subdifferentials of Lipschitz functions, *Trans. Amer. Math. Soc.* **264** (1981), 77-89.
- [23] VIAL, J.-P., Strong and weak convexity of sets and functions, *Math. Oper. Res.* **8** (1983), 231-259.

Aris DANIILIDIS

Departament de Matemàtiques
 Universitat Autònoma de Barcelona
 E-08193 Bellaterra (Cerdanyola del Vallès), Spain.
 e-mail: arisd@mat.uab.es
<http://mat.uab.es/~arisd>

Jérôme MALICK

INRIA, Rhone-Alpes
 655 avenue de l'Europe
 Montbonnot, St Martin
 F-38334 Saint Ismier, France
 email: jerome.malick@inria.fr
<http://www.inrialpes.fr/bipop/people/malick/>